PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/00 | A2

(11) International Publication Number: WO 99/49024

(43) International Publication Date: 30 September 1999 (30.09.99)

(21) International Application Number: PCT/US99/06506

(22) International Filing Date: 24 March 1999 (24.03.99)

(30) Priority Data:
60/079,249 25 March 1998 (25.03.98) US

(71) Applicant *(for all designated States except US)*: PLANET BIOTECHNOLOGY, INC. [US/US]; 2462 Wyandotte Street, Mountain View, CA 94045 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: WYCOFF, Keith, L. [–/US]; 2399 Carmel Drive, Palo Alto, CA 94303 (US). JAISWAL, Sudhir, K. [–/US]; Apartment 285, 100 N. Whisman Road, Mountain View, CA 94043 (US).

(74) Agents: CIOTTI, Thomas, E. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304–1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHODS AND COMPOSITIONS FOR PRODUCTION OF MULTIMERIC PROTEINS IN TRANSGENIC PLANTS

(57) Abstract

This invention is related to the field of methods and compositions for production of multimeric proteins in transgenic plants. The invention provides a method for producing a heterologous multimeric protein in a transgenic plant cell by transforming a plant cell with a plurality of naked plasmids. Each plasmid encodes less than all of the polypeptide components of the multimeric protein, and the plurality of transformed plasmids encodes all of the polypeptide components of the multimeric protein. The invention also provides a transgenic plant or plant cell expressing a multimeric protein that is heterologous to the plant cell.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHODS AND COMPOSITIONS FOR PRODUCTION OF MULTIMERIC PROTEINS IN TRANSGENIC PLANTS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the priority benefit of provisional U.S. Patent Application 60/079,249, filed March 25, 1998, pending, which is hereby incorporated herein by reference in its entirety.

TECHNICAL FIELD

This invention is related to the field of methods and compositions for production of multimeric proteins in transgenic plants.

BACKGROUND

Transgenic plants have been studied over the past several years for potential use in low cost production of high quality, biologically active mammalian proteins. Of the various mammalian proteins studied to date, monoclonal antibodies have received the most attention because of their potential value as therapeutic and clinical reagents (reviewed in Ma and Hein, 1995; Ma and Hein, 1996). Initially, individual transgenic tobacco plants each expressing either the heavy or light immunoglobulin chains were crossed, generating a hybrid plant that co-expressed both chains. More recently, through successive crosses, it has been possible to generate tobacco plants expressing functional secretory IgA (SIgA) consisting of heavy and light chains plus a J chain and secretory component. However, such cross-breeding is a time-consuming process and in some plants, like alfalfa, that do not readily self-fertilize, it is impractical. There thus remains a need for methods and compositions for simultaneous introduction of genes into plants encoding multiple protein chains capable of associating

together or assembling together, for example antibody chains, and rapid scale-up to commercial production levels without laborious crosses to commercial varieties.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for the transformation of plant cells with multiple genes, and proper association or assembly of multimeric proteins that are heterologous to the plant cells. At least two separate plasmid constructs encoding the individual polypeptide components of the multimeric protein are used. Preferably, the plasmids are introduced into the target cells as naked DNA. Typically, these plasmids used in the invention are made as "expression cassettes" which include the required regulatory sequences. For example, such DNA constructs (expression cassettes) may include: a promoter that is functional in a given host plant cell; nucleic acid encoding a signal peptide fused to nucleic acid molecule encoding a subunit protein; nucleic acid encoding the amino acid sequence KDEL for retention in the endoplasmic reticulum; and a polyadenylation signal sequence. The transformed plant cells are cultured under conditions to produce the assembled protein which can then be isolated.

In one aspect, the multimeric protein comprises an immunoglobulin molecule, a receptor ligand complex, a homodimer or heterodimer. Typically, the multimeric protein is biologically active. The present invention encompasses use of plant cells from a dicot genus, for example *Nicotiana*, or from a monocot genus, for example, *Lemna*.

In another aspect, additional sequences are introduced into the plant cells. These additional sequences include, but are not limited to, sequences encoding the amino acid sequence KDEL (which aids in accumulating the multimer in the endoplasmic reticulum); a signal sequence (which facilitates the targeting of the multimer to the endoplasmic reticulum); a selectable marker (which allows for identification of transformed cells) and a leader sequence. Thus, also

encompassed by the present inventions are DNA constructs for high level expression and secretion of heterologous proteins in plant cells. These additional sequences may be included on the plasmids carrying one or more structural components of the multimer. Alternatively, these additional sequences can be introduced on additional plasmids.

In yet another aspect, the present invention includes microparticles coated with multiple plasmids, where at least two of these multiple plasmids encode structural components made of a multimeric protein. Microparticles are preferably of inert metals such as tungsten or gold and are preferably between about 0.5 and 1.5 microns in size.

The present invention also includes transgenic plants and cells that produce multimeric proteins. These plants and plant cells are characterized by adjacent integration of multiple expression cassettes, wherein each expression cassette encodes at least one but less than all of the polypeptide components of a multimeric protein.

As will become apparent, preferred features and characteristics of one aspect of the invention are applicable to any other aspect of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic depicting the plasmid SSpHuK (pSSpHuK).

Figure 2 is a schematic depicting the plasmid SSpHuA2 (pSSpHuA2).

Figure 3 is a schematic depicting the plasmid SHuJ (pSSpSHuJ).

Figure 4 is a schematic depicting the plasmid SSpHuSC (pSSpHuSC).

Figure 5 (SEQ ID NOS:1 and 2) depicts the nucleotide and amino acid sequence of the coding region of pSSpGuyHuK.

Figure 6 (SEQ ID NOS: 3 and 4) depicts the nucleotide and amino acid sequence of the coding region of pSSpGuyHuA2.

Figure 7 (SEQ ID NOS: 5 and 6) depicts the nucleotide and amino acid sequence of the coding region of pSHuJ.

Figure 8 (SEQ ID NOS: 7 and 8) depicts the nucleotide and amino acid sequence of the coding region of pSHuSC.

MODES FOR CARRYING OUT THE INVENTION

Throughout this disclosure, various publications, patents and published patent specifications are referenced by an identifying citation. The disclosures of these publications, patents and published patent specifications are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

Definitions

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of immunology, molecular biology, microbiology, cell biology and recombinant DNA, which are within the skill of the art. See, *e.g.*, Sambrook, Fritsch and Maniatis, MOLECULAR CLONING: A LABORATORY MANUAL, 2nd edition (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (F. M. Ausubel, et al. eds., (1987)); the series METHODS IN ENZYMOLOGY (Academic Press, Inc.): PCR 2: A PRACTICAL APPROACH (M.J. MacPherson, B.D. Hames and G.R. Taylor eds. (1995)), Harlow and Lane, eds. (1988) ANTIBODIES, A LABORATORY MANUAL, and METHODS IN MOLECULAR BIOLOGY vol. 49, "PLANT GENE TRANSFER AND EXPRESSION PROTOCOLS", H. Jones, eds. (1995).

As used in the specification and claims, the singular form "a", "an" and "the" include plural references unless the context clearly dictates otherwise. For example, the term "a cell" includes a plurality of cells, including mixtures thereof.

A "variable region" of an antibody refers to the variable region of the antibody's light chain or the variable region of the heavy chain, either alone or in combination.

As used herein, a "polynucleotide" is a polymeric form of nucleotides of any length, which contain deoxyribonucleotides, ribonucleotides, and/or their analogs. The terms "polynucleotide" and "nucleotide" as used herein are used interchangeably. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The term "polynucleotide"

includes double- , single-stranded, and triple-helical molecules. Unless otherwise specified or required, any embodiment of the invention described herein that is a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double stranded form.

The term "polypeptide" is used in its broadest sense to refer to a compound of two or more subunit amino acids, amino acid analogs, or peptidomimetics. The subunits may be linked by peptide bonds. In another embodiment, the subunit may be linked by other bonds, e.g. ester, ether, etc. As used herein the term "amino acid" refers to either natural and/or unnatural or synthetic amino acids, including glycine and both the D or L optical isomers, and amino acid analogs and peptidomimetics. A peptide of three or more amino acids is commonly called an oligopeptide if the peptide chain is short. If the peptide chain is long, the peptide is commonly called a polypeptide or a protein.

A "multimeric protein" as used herein refers to a globular protein containing more than one separate polypeptide or protein chain associated with each other to form a single globular protein *in vitro* or *in vivo*. The multimeric protein may consist of more than one polypeptide of the same kind to form a homodimeric or homotrimeric protein; the multimeric protein may also be composed of more than one polypeptide having distinct sequences to form, e.g., a heterdimer or a heterotrimer. Non-limiting examples of multimeric proteins include immunoglobulin molecules, receptor dimer complexes, trimeric G-proteins, and any enzyme complexes.

An "immunoglobulin molecule" or "antibody" is a multimeric protein containing the immunologically active portions of an immunoglobulin heavy chain and immunoglobulin light chain covalently coupled together and capable of specifically combining with antigen. An "immunoglobulin combining site" or "binding domain" is formed from the folding of variable domains of an antibody molecule(s) to form three-dimensional binding spaces with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows an immunological reaction with the antigen. An antibody

combining site may be formed from a heavy and/or a light chain domain (VH and VL, respectively), which form hypervariable loops which contribute to antigen binding. The immunoglobulins, or antibody molecules, are a large family of molecules that include several types of molecules, such as IgD, IgG, IgA, IgM and IgE. The term "immunoglobulin molecule" includes, for example, hybrid antibodies, or altered antibodies, and fragments thereof, including but not limited to Fab fragment(s), and Fv fragment.

An Fab fragment of an immunoglobulin molecule is a multimeric protein consisting of the portion of an immunoglobulin molecule containing the immunologically active portions of an immunoglobulin heavy chain and an immunoglobulin light chain covalently coupled together and capable of specifically combining with an antigen. Fab fragments can be prepared by proteolytic digestion of substantially intact immunoglobulin molecules with papain using methods that are well known in the art. However, a Fab fragment may also be prepared by expressing in a suitable host cell the desired portions of immunoglobulin heavy chain and immunoglobulin light chain using methods disclosed herein or any other methods known in the art.

An Fv fragment of an immunoglobulin molecule is a multimeric protein consisting of the immunologically active portions of an immunoglobulin heavy chain variable region and an immunoglobulin light chain variable region covalently coupled together and capable of specifically combining with an antigen. Fv fragments are typically prepared by expressing in suitable host cell the desired portions of immunoglobulin heavy chain variable region and immunoglobulin light chain variable region using methods described herein and/or other methods known to artisans in the field.

"Heterologous" means derived from a genotypically distinct entity from that of the rest of the entity to which it is compared. For example, a polynucleotide introduced by genetic engineering techniques into a different cell is a heterlogous polynucleotide (and, when expressed, can encode a heterologous polypeptide). In particular, the term "heterologous" as applied to a multimeric protein means that the multimer is expressed in a host cell that is genotypically

distinct from the host cell in which the multimer is normally expressed. For example, the exemplified human IgA multimeric protein is heterologous to a plant cell.

The term "biologically active", as used herein, refers to a multimer having structural, regulatory, or biochemical functions of a naturally occurring molecule expressed in its native host cell. For instance, a biologically active immunoglobulin produced in a plant cell by the methods of this invention has the structural characteristics of the naturally occurring molecule, and/or exhibits antigen binding specificity of the naturally occurring antibody that is present in the host cell in which the molecule is normally expressed.

A "gene" refers to a polynucleotide containing at least one open reading frame that is capable of encoding a particular protein after being transcribed and translated.

As used herein, "expression" refers to the process by which polynucleotides are transcribed into mRNA and/or the process by which the transcribed mRNA is subsequently being translated into peptides, polypeptides, or proteins. If the polynucleotide is derived from genomic DNA, expression may include splicing of the mRNA in a eukaryotic cell.

The term "construct" refers to an artificially assembled DNA segment to be transferred into a target plant tissue or cell. Typically, the construct will include the gene of a particular interest, a marker gene and appropriate control sequences. The term "plasmid" refers to an autonomous, self-replicating extrachromosomal DNA molecule. In a preferred embodiment, the plasmid constructs of the present invention contain sequences coding for heavy or light chain constant regions of an antibody. The plasmids also preferably contain sequences encoding a signal peptide, a leader sequence, a sequence for retention in the endoplasmic reticulum (ER), such as KDEL. Plasmid constructs containing suitable regulatory elements are also referred to as "expression cassettes." In a preferred embodiment, a plasmid construct can also contain a screening or selectable marker, for example an antibiotic resistance gene. The term "naked

DNA" or "naked plasmid" refers to plasmids which are not packaged in a gene delivery vehicle, for example agrobacterium.

The terms "screening marker", "selectable marker" and "reporter gene" are used to refer to a gene that encodes a product that can readily be assayed. For example, reporter genes can be used to determine whether a particular DNA construct has been successfully introduced into a cell, organ or tissue. Non-limiting examples of selectable markers include genes encoding for antibiotic resistance, *e.g.*, ampicillin, kanamycin or the like. Other selection markers will be known to those of skill in the art.

A "signal sequence" is a nucleic acid sequence encoding the "signal peptide" located typically at the N-terminus of a polypeptide, which direct the polypeptide to a specific cell compartment, e.g. the endoplasmic reticulum, wherein the assembly of multimer takes place.

A "primer" is a short polynucleotide, generally with a free 3' -OH group, that binds to a target or "template" potentially present in a sample of interest by hybridizing with the target, and thereafter promoting polymerization of a polynucleotide complementary to the target. A "polymerase chain reaction" ("PCR") is a reaction in which replicate copies are made of a target polynucleotide using a "pair of primers" or a "set of primers" consisting of an "upstream" and a "downstream" primer, and a catalyst of polymerization, such as a DNA polymerase, and typically a thermally-stable polymerase enzyme. Methods for PCR are well known in the art, and taught, for example in MacPherson, et al., PCR: A Practical Approach (IRL Press at Oxford University Press (1991)). All processes of producing replicate copies of a polynucleotide, such as PCR or gene cloning, are collectively referred to herein as "replication." A primer can also be used as a probe in hybridization reactions, such as Southern or Northern blot analyses.

"Hybridization" refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogstein binding, or in any other sequence-specific manner. The

complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

When hybridization occurs in an antiparallel configuration between two single-stranded polynucleotides, the reaction is called "annealing" and those polynucleotides are described as "complementary". A double-stranded polynucleotide can be "complementary" or "homologous" to another polynucleotide, if hybridization can occur between one of the strands of the first polynucleotide and the second. "Complementarity" or "homology" (the degree that one polynucleotide is complementary with another) is quantifiable in terms of the proportion of bases in opposing strands that are expected to form hydrogen bonding with each other, according to generally accepted base-pairing rules.

A "transgenic plant" refers to a genetically engineered plant or progeny of genetically engineered plants. The transgenic plant usually contains material from at least one unrelated organism, such as from a virus, another plant, or animal.

As used herein, the term "isolated" means separated from constituents, cellular and otherwise, in which the polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof, are normally associated with in nature. As is apparent to those of skill in the art, a non-naturally occurring polynucleotide, peptide, polypeptide, protein, antibody, or antibody fragments, does not require "isolation" to distinguish it from its naturally occurring counterpart. In addition, a "concentrated", "separated" or "diluted" polynucleotide, peptide, polypeptide, protein, antibody, or antibody fragments, is distinguishable from its naturally occurring counterpart in that the concentration or number of molecules per volume is greater than "concentrated" or less than "separated" than that of its naturally occurring counterpart. A polynucleotide, peptide, polypeptide, protein, antibody, or fragments thereof , which differs from the naturally occurring counterpart in its primary sequence or for example, by its glycosylation pattern, need not be present in its isolated form since it is distinguishable from its naturally

occurring counterpart by its primary sequence, or alternatively, by another characteristic such as glycosylation pattern. A protein produced in a plant cell is provided as a separate embodiment from the naturally occurring protein isolated from another type of cell in which it is produced in nature.

A "control" is an alternative subject or sample used in an experiment for comparison purpose. A control can be "positive" or "negative". For example, where the purpose of the experiment is to determine the presence of an exogenously introduced plasmid or the expression of a polypeptide encoded by such plasmid in a plant transformant or its progenies, it is generally preferable to use a positive control (a plant or a sample from a plant, carrying such plasmid and/or expressing the encoded protein), and a negative control (a plant or a sample from a plant lacking the plasmid of interest and/or expression of the polypeptide encoded by the plasmid).

The present invention provides methods for producing multimeric proteins by introducing multiple plasmids into a plant cell. The inventors have shown that multimeric proteins are assembled and accumulated when the individual polypeptide components are encoded by separate plasmids. In addition, constructs carrying sequences encoding signal peptides, leader sequences, selectable marker sequences and accumulation sequences have also been shown to transform plant cells and aid in the production of functional multimeric proteins. The invention, therefore, provides a fast and efficient method of making heterologous multimeric proteins in plant cells. The plants transformed by the methods described herein are also stably integrated and their progeny also produce the multimeric protein(s).

The methods described herein also result in a transgenic plant that is genetically distinguishable from plants which have been cross-bred to produce multimers. Similarly, the plants and plant cells described herein are also distinguishable from plants into which a single plasmid carrrying sequences encoding components of a multimeric protein has been introduced. In particular, the transgenic plants described herein are characterized in that the separate

expression cassettes of the multiple plasmids are integrated adjacent to one another in the host genome. In contrast, single plasmid tranformants will carry a single integrated expression cassette, while cross-bred plants will carry randomly integrated (non-adjacent) expression cassettes. These differences can be determined by methods known in the art, for example, by Southern blotting.

Transformation of Plant Cells

The present invention provides for methods of producing multimeric proteins by transforming plant cells with a plurality of plasmids. Typically, the plasmids are introduced as naked DNA, for example by particle bombardment. Sequence coding for individual components of the multimer is found on at least two plasmids. In one embodiment, each component is encoded for by a different plasmid. In another embodiment, for example where the multimer is made up of more than two structural components, it is possible to use a plasmid encoding all but one of the structural components in conjunction with a plasmid encoding the missing component. Non-limiting examples of multimeric proteins and their respective structural components include immunoglobulins made up of heavy and light chains and optional J-chains and secretory components; receptor molecules made up of two or more components and the like. In addition, using the methods described herein, it is possible to assemble a functional multimeric protein using plasmids carrying fragments of the components of the multimer, *e.g.*, Fab or Fv fragments of immunoglobulins.

Methods of making plasmid constructs are well known in the art and described for example, in Sambrook, *supra.* Plasmids will typically contain at least one structural component and appropriate regulatory sequences, such as promoters suitable for use in plants. Non-limiting examples of promoters include constitutive (or nearly constitutive) promoters from agrobacterium such as a superpromoter; nopaline synthase; octopine synthase; mannopine synthase and 1'2'; the rbcS (small subunit of ribulose bis-phosphate carboxylase promoter from plants; chlorella virus adenine methyl transferase (AMT) and ubiquitin. Examples of inducible promoters include, chitinase (from bean), chalcone synthase

phenylalanine ammonia lyase and HRGP. Tissue-specific promoters may also be used, for example, legumin (or other seed storage protein promoters), patatin and the like. For a discussion of the 35S promoter, see, for example, Odell JT et al. (1985) *Nature* 313: 810-812 and U.S. Patent No. 5,352,605, issued October 1994).

At least two of the plasmids will each encode at least one of the structural components of the multimer. In a preferred embodiment, the structural components are linked to signal peptides necessary for the assembly of a multimeric protein. Suitable signal peptides can be readily determined by those of skill in the art and include, by way of example, the signal peptide of 2S2 storage protein of A. thaliana, Vicia faba legumin B4, alpha amylase or patatin. One of skill in the art could readily construct a plasmid carrying sequence encoding a polypeptide in proper frame with a signal peptide coding sequence.

In another preferred embodiment, at least one of the plasmids used in transformation will encode a peptide, for example the amino acid sequence "KDEL," which results in accumulation of the multimer in the endoplasmic reticulum. In yet another preferred embodiment, at least one of the plasmids used in transformation encodes a selectable marker. Suitable selectable markers for plants will be known to those in the art and will be appropriate for use in transgenic plants, which are regenerated via direct shoot organogenesis or somatic embryogenesis. Cloning and expression vectors typically contain a selectable marker (for example, a gene encoding a protein necessary for the survival or growth of a host cell transformed with the vector). Only those host cells into which a selectable gene has been introduced will grow under selective conditions. Typical selection genes either: (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, kanamycin, methotrexate; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media. The choice of the proper marker gene will depend on the host cell, and appropriate genes for different hosts are known in the art.

In one embodiment, described in the Examples below, generic expression vectors to fuse any mouse variable region to human heavy and light chain constant regions for expression in plants were used. The heavy chain vector,

pSSpHuA2 (Figure 2; Figure 6; SEQ ID NOS: 3 and 4), contains sequence encoding a bean legumin signal peptide and the human IgA2m(2) constant region with Spe I and Sac I sites in between for cloning variable regions. The light chain vector, pSSpHuK (Figure 1; Figure 5: SEQ ID NOS: 1 and 2), contains sequence encoding a bean legumin signal peptide and the human kappa constant region with Spe I and Hind III sites in between for cloning variable regions. Expression of these genes is under the control of the same promoter and terminator as in the heavy chain vector. Heavy and light chain constructs containing the variable regions of an anti-*S. mutans* SA I/II antibody (Guy's 13; see Figs. 1 and 2), along with constructs for the expression of human J chain, pSHuJ (Figure 3; Figure 7; SEQ ID NOS: 5 and 6) and secretory component, pSHuSC (Figure 4; Figure 8; SEQ ID NOS: 7 and 8), are used for transformation (USSN 08/430,000 incorporated by reference).

The target plant cells may be in the form of whole plants, isolated cells or protoplasts. Preferably, the cells are "intact" in that the cell comprises an outer layer of cell wall, typically composed of cellulose for protection and maintaining the rigidity of the plant cell. These plant cells may be transformed with the plurality of plasmids using any method known in the art. In a particularly preferred embodiment, the plant cells are bombarded with naked DNA using microparticles coated with the plurality of plasmids. Bombardment with DNA-coated microprojectiles has been successfully used to produce stable transformants in both plants and animals (see, for example, Sanford *et al.*, 1993, *infra*). As many as 12 different plasmids, introduced simultaneously, have been stably integrated into soybean cells via particle bombardment (Hadi *et al.*, 1996, *infra*). However, these plasmids did not encode polypeptides and expression was not shown. In rice, as many as 13 unrelated genes were inserted into the genome, expressed, and stably inherited (Chen *et al.*, 1998, *infra*). However, assembly of polypeptides expressed from multiple plasmids into functional mulitimeric protein complexes in regenerated plants has not previously been demonstrated.

Suitable micorparticles for bombardment are available for example, from BioRad (*e.g.*, Bio-Rad's PDS-1000/He). Typically, particles suitable for use in

the methods described herein are made of metal, preferably tungsten or gold. Their average diameter is generally in the range of about 0.5 microns to about 1.5 microns.

As noted above, the methods described herein involve the use of at least two plasmids each encoding less than all of the structural components of a multimer. In addition, these at least two plasmids, or additional plasmids, may encode signal peptides, leader sequences, selectable marker genes or the like. Although not necessarily required, it is believed that the microparticles should be coated with all the plasmids to be used. Without being bound by one theory, it is believed to be unlikely that more than one microparticle will enter and stably transform a plant cell. Accordingly, it is preferable to ensure that each microparticle carries the plasmids of interest.

As will be evident to those of skill in the art, the particle bombardment protocol can be optimized for any plant by varying parameters such as He pressure, quantity of coated particles, distance between the macrocarrier and the stopping screen and flying distance from the stopping screen to the target. An optimal protocol is described in the Examples.

In one embodiment, secretory IgA molecules can be assembled by using plasmid containing heavy chain, light chain, J chain and secretory component, are used for particle bombardment of leaf disks. A fifth vector containing an antibiotic resistance gene, such as pSZeo (containing the Phe^r gene encoding resistance to the antibiotic Zeocin; Perez et al., 1989; Drocourt et al, 1990) or pBMSP-1 (containing the Kan^r gene encoding resistance to kanamycin; S. Gelvin, Purdue University) is also included. The transgenic plants obtained may be screened for production of assembled SIgA using the methods described below.

Selection of Plant Transformants

Transformation of a host plant with a plurality of plasmids, each plasmid encoding less than all of the components of the multimeric protein, yields a population of transformants that may carry some or all of the plasmids of interest. To select those transformants containing multiple plasmids encoding each and

every component of the multimeric protein, one generally proceeds by detecting the presence of the plasmids, and/or the expression of each component of the multimer in the transformants or the transgenic plant produced therefrom.

The presence of all plasmids encoding each component of the multimeric protein can be confirmed by hybridization assays, amplification reactions using a probe or a primer pair derived from the plasmids used for transformation. Preferably, the probe or primer pair comprises the sequences complementary to the nucleic acid encoding the full-length or a fragment of each polypeptide chain that constitutes the multimer. In an alternative, the success of transformation can also be determined by restriction enzyme digestion, in which the appearance of restricted fragments representative of the plasmids is indicative of the presence of the exogenously introduced plasmids.

The test nucleic acid can be obtained from the cultured plant transformants, such as the plasmid-bombarded leaf disks, or the regenerated plant therefrom. The nucleic acid to be tested can be extracted from plant according to standard methods in the art. For instance, nucleic acid can be isolated using various lytic enzymes or chemical solutions according to the procedures set forth in Sambrook et al. ("Molecular Cloning: A Laboratory Manual", Second Edition,1989), or extracted by nucleic-acid-binding resins following the accompanying instructions provided by manufactures.

Hybridization can be performed under conditions of different "stringency". Conditions that vary levels of stringency are well known in the art (see, for example, Sambrook, et al. supra). Briefly, relevant conditions include temperature, ionic strength, time of incubation, the presence of additional solutes in the reaction mixture such as formamide, and the washing procedure. Higher stringency conditions are those conditions, such as higher temperature and lower sodium ion concentration, which require higher minimum complementarity between hybridizing elements for a stable hybridization complex to form. In general, a low stringency hybridization reaction is carried out at about 40 ^{0}C in 10 x SSC or a solution of equivalent ionic strength/temperature. A moderate stringency hybridization is typically performed at about 50 ^{0}C in 6 X SSC, and a

high stringency hybridization reaction is generally performed at about 60 ^{0}C in 1 X SSC. In determining the presence of an exogenously introduced plasmid carrying a gene of interest in the plant transformants, stringent hybridization is preferred.

For the purpose of this invention, amplification means any method employing a primer-dependent polymerase capable of replicating a target sequence with reasonable fidelity. Amplification may be carried out by natural or recombinant DNA-polymerases such as T7 DNA polymerase, Klenow fragment of E. coli DNA polymerase, and reverse transcriptase.

A preferred amplification method is PCR. General procedures for PCR are taught in "PCR: a practical approach" (M. MacPherson et al., IRL Press at Oxford University Press 1991). After amplification, the resulting DNA fragments can be detected by agarose gel electrophoresis followed by visualization with ethidium bromide staining and ultraviolet illumination. A specific amplification of the structural gene can be verified by demonstrating that the amplified DNA fragment has the predicted size, exhibits the predicated restriction digestion pattern, and/or hybridizes to the correct cloned DNA sequence.

The probe or primers employed in the hybridization and amplification reactions can be conjugated to a detectable marker, e.g., an enzymatic label or a radioisotope for detection of nucleic acid and/or expression of the gene in a cell. A wide variety of appropriate detectable markers are known in the art, including fluorescent, radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of giving a detectable signal. In preferred embodiments, one will likely desire to employ a fluorescent label or an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of radioactive or other environmental undesirable reagents. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide a means visible to the human eye or spectrophotometrically, to identify specific hybridization with complementary nucleic acid-containing samples.

To ascertain the expression of each polypeptide contained in the multimeric protein, plant transformants or progeny made therefrom can be tested

in an immunoassay using antibodies specific for individual polypeptides that constitute the multimer. Procedures for carrying out immunoassays which include but are not limited to immunoblotting and immunoprecipitation which are well established and commonly practiced by an ordinary skill in the art. Briefly, the reaction is performed by contacting the detecting antibody with a sample containing the test proteins of the plant transformants or their progenies under conditions that will allow a complex to form between the antibody and the target polypeptide. The sample of test proteins can be prepared by homogenizing the plant transformants or their progenies made therefrom, and optionally solubilizing the test protein using detergents, preferably non-reducing detergents such as triton and digitonin. The binding reaction in which the test proteins are allowed to interact with the detecting antibodies may be performed in solution, or on a solid tissue sample, for example, using tissue sections or solid support that has been immobilized with the test proteins. The formation of the complex can be detected by a number of techniques known in the art. For example, the antibodies may be supplied with a label and unreacted antibodies may be removed from the complex; the amount of remaining label thereby indicating the amount of complex formed.

The amount of the target polypeptides that are immunologically reactive with the detecting antibodies can also be quantified by standard quantitative immunoassays. For instance, the target protein may be solubilized or extracted from a solid tissue sample, and mixed with a pre-determined non-limiting amount of the reagent antibody specific for the protein. The detecting antibody may contain a directly attached label, such as an enzyme or a radioisotope, or a second labeled reagent may be added, such as protein A. For a solid-phase assay, unreacted reagents are removed by washing. For a liquid-phase assay, unreacted reagents are removed by some other separation technique, such as filtration or chromatography. The amount of label captured in the complex is positively related to the amount of target protein present in the test sample. Alternatively, a competitive assay in which the target protein is tested for its ability to compete with a labeled analog for binding sites on the specific antibody. In this case, the amount of label captured is negatively related to the amount of target protein

present in a test sample. Results obtained using any such assay on a sample from a plant transformant or a progeny thereof is compared with those from a non-transformed source as a control.

Detection of the Assembled, Biologically Active Multimeric Proteins

Production of biologically active multimer in a plant transformant or its progeny generally requires assembly of individual polypeptide components to assume a proper conformation that confers a biological activity. Selected plant transformants and/or the progenies made therefrom can be assayed for the presence of the assembled, biologically active multimeric protein using a variety of methods well known in the art. Such methods include ELISA, Western blotting, immunoprecipitation, and any assay designed to detect a functional multimeric protein by measuring, e.g. its enzymatic activity or immunological activity. The enzymatic activity of a multimer include but are not limited to phosphorylation, glycosylation, peptidase activity, GTPase activity, voltage-dependent ATPase activity, and any combinations thereof.

Preferred screening assays are those where the biological activity of an assembled multimeric protein is detected in such a way as to produce a detectable signal. This signal may be produced directly or indirectly and such signals include, for example, the production of a complex via protein-protein interaction, induction of a conformational change of the multimer, formation of a catalytic reaction product, the release or uptake of energy, and the like. Where the multimer is an immunglobulin molecule, its biological activity can be assessed by its ability to bind specifically to its cognate antigen. As used herein, the specificity of an antibody refers to the ability of the antibody to distinguish cognate (or native) antigens from any other unrelated antigens. Where the multimer is a receptor-ligand complex or a complex of receptor molecules, its biological activity can be confirmed by the ability of the complex to interact with downstream signaling molecules and/or its ability to phosphorylate substrates on the tyrosine, serine and/or histidine residues. The biological activity of a trimeric G protein can be ascertained by the presence of GTPase activity. Assaying

procedures for phosphorylation and GTPase activity, as well as protocols for detecting protein-protein interactions are well established in the art, and thus are not detailed herein.

Isolation of Biologically Active Multimeric Proteins

A transgenic plant of the present invention expressing the desired multimeric protein is cultivated to isolate the desired multimeric protein they contain. After cultivation, the transgenic plant is harvested to recover the produced multimeric protein. This harvesting step may consist of harvesting the entire plant, or only the leaves, or roots of the plant. This step may either kill the plant, or if only the portion of the transgenic plant is harvested, may allow the remainder of the plant to continue to grow. In a specific embodiment, harvesting the transgenic plant may further comprise (a) homogenizing at least a portion of said transgenic plant to produce a plant pulp; (b) extracting the multimeric protein from the plant pulp to produce a multimeric protein containing solution; and (c) isolating said multimeric protein from said solution. More specifically, at least a portion of the transgenic plant is homogenized to produce a plant pulp using methods well known to one skilled in the art. This homogenization may be done manually, by a machine, or by a chemical means as long as the transgenic plant portions are broken up into small pieces to produce a plant pulp. This plant pulp consists of a mixture of varying sizes of transgenic plant particles. The size of the plant particles and the amount of variation in size that can be tolerated will depend on the exact method used to extract the multimeric protein from the plant pulp and these parameters are well known to one skilled in the art. The multimeric protein is extracted from the plant pulp produced above to form a multimeric protein containing solution. Such extraction processes are common and well known to articians in this art. For example, the extracting step may consist of soaking or immersing the plant pulp in a suitable solvent. A suitable solvent dissolves the multimeric protein present in the plant pulp to produce a multimeric protein containing solution. Solvents useful for such an extraction process include but are not limited to aqueous solvents, organic solvents and combinations of both. A

preferred solvent is non-reducing detergent such as digitonin or Triton-X100. The multimeric protein can then be isolated from the solution produced above by methods such as immunoprecipitation with antibody, and standard techniques such as gel filtration, ion-exchange, reversed-phase, and affinity chromatography. For such methodology, see for example Deutscher et al. (1999) GUIDE TO PROTEIN PURIFICATION: METHODS IN ENZYMOLOGY (Vol. 182, Academic Press).

REFERENCES

Bäumlein H, Wobus U, Pustell J, Kafatos FC. 1986. The legumin gene family: structure of a B type gene of *Vicia faba* and a possible legumin gene specific regulatory element. *Nucl. Acids Res.* 14: 2707-2713.

Chen L, Marmey P, Taylor NJ, Brizard J-P, Espinoza C, D'Cruz P, Huet H, Zhang S, de Kochko A, Beachy RN, Fauquet CM. 1998. Expression and inheritance of multiple transgenes in rice plants. *Nat. Biotechnol.* 16: 1060-1064.

Chintalacharuvu K.R., Raines M., Morrison S.L. 1994. Divergence of human alpha-chain constant region gene sequences. A novel recombinant alpha 2 gene. *Journal of Immunology* 152: 5299-5304.

Depicker A, Stachel S, Dhaese P, Zambryski P, Goodman HM. 1982. Nopaline synthase: transcript mapping and DNA sequence. *J. Mol. Appl. Genet.* 1: 561-573

Drocourt D., Calmels T., Reynes J.-P., Baron M., and Tiraby G. 1990. Cassettes of the *Streptoalloteichus hindustanus ble* gene for transformation of lower and higher eukaryotes to phleomycin resistance. *Nucleic Acids Research*, 18:4009.

Hadi M.Z., McMullen M.D., and Finer J.J. 1996. Transformation of 12 different plasmids into soybean via particle bombardment. *Plant Cell Reports*, 15:500-505.

Hieter P.A., Max E.E., Seidman J.G., Maizel J.V.Jr., Leder P. 1980 Cloned human and mouse kappa immunoglobulin constant and J region regions conserve homology in functional segments. Cell 22: 197-207

Horsch R.B., Fry J.E., Hoffmann N.L., Eichholtz D., Rogers S.G., and Fraley R.T. 1985. A simple and general method for transferring genes into plants. *Science*, 227:1229-1231.

Kozak M. 1986. Point mutations define a sequence flaking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell*, 44: 283-292.

Krajci P, Solberg R, Sandberg M, Oyen O, Jahnsen T, Bradtzaeg P. 1989. Molecular cloning of the human transmembrane secretory component (poly-Ig receptor) and its mRNA expression in human tissues. *Biochem. Biophys. Res. Commun.* 158: 783

Ma J.K.-C., and Hein M. 1995. Immunotherapeutic potential of antibodies produced in plants. *TIBTECH*, 13:522-527.

Ma J.K.-C., and Hein M.B. 1996. Antibody production and engineering in plants. In Annals of the New York Academy of Sciences: Engineering plants for commercial products and applications. Edited by G. B. Collins and R. J. Shepherd. 72-81. New York, NY: The New York Academy of Sciences.

Max E., Korsmeyer S. 1985. Human J Chain Gene. *Journal of Experimental Medicine* 161: 832-849.

Ni M, Cui D, Einstein J, Narasimhulu S, Vergara CE, Gelvin SB (1995) Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes. *Plant Journal* 7: 661-676

Perez P., Tiraby G., Kallerhoff J., and Perret J. 1989. Phleomycin resistance as a dominant selectable marker for plant cell transformation. *Plant Molecular Biology*, 13:365-373.

Porath, D. Dec. 14, 1993. Aquaculture production of duckweed for human consumption. United States Patent No. 5,269,819.

Sanford J.C., Smith F.D., and Russell J.A. 1993. Optimizing the biolistic process for different biological applications. *Methods in Enzymology*, 217:483-509.

Smith R, Lehner T.1989. Characterisation of monoclonal antibodies to common protein epitopes on the cell surface of *Streptococcus mutans* and *Streptococcus sobrinus*. *Oral Microbiology and Immunology* 4: 153-158.

EXAMPLES

Example 1 Transformation of Tobacco with SIgA

Construction of Secretory IgA Expression Vectors

We have constructed generic expression vectors to fuse any mouse variable region to human heavy and light chain constant regions for expression in plants. The heavy chain vector is called pSSpHuA2 (Figure 2), and contains sequence encoding a bean legumin signal peptide and the human IgA2m(2) constant region with *Spe* I and *Sac* I sites in between for cloning variable regions. The light chain vector is called pSSpHuK(Figure 1), and contains sequence encoding a bean legumin signal peptide and the human kappa constant region with Spe I and Hind III sites in between for cloning variable regions. Expression is under the control of the same promoter and terminator as in the heavy chain vector. Heavy and light chain constructs containing the variable regions of an anti-*S. mutans* SA I/II antibody (Guy's 13; see Figs. 1 and 2), along with constructs for the expression of human J chain (pSHuJ, Fig. 3) and secretory component (pSHuSC, Fig. 4), were used for transformation.

The expression vectors constructed for plant transformation were:

1. The plasmid pSSpGuyHuK is 4163 base pairs in length. Nucleotides 16-1139 represent the Superpromoter (Ni et al., 1995), which drives expression in plant cells. Nucleotides 1140-1864 encode the light chain (Figure 5) and comprise a sequence encoding a mouse light chain variable/human kappa constant hybrid with linker sequences. A consensus Kozak sequence (Kozak, 1986) is included (nt 1147-1154) to enhance translation initiation, and the native mouse signal peptide has been replaced with the signal peptide from bean legumin (nt 1150-1218; Bäumlein et al., 1986). The variable region (nt 1219-1533) is from the murine monoclonal Guy's 13 (Smith and Lehner, 1989, US Patent No. 5,518,721 and 5,352,446). The sequence of the human kappa constant region (nt 1534-1864) has been previously published (Hieter et al., 1980). Nucleotides 1865-2118 derive from the nopaline synthase 3' end (transcription termination

and polyadenlyation signal; Depicker et al., 1982). The remainder of the plasmid derives from the vector pSP72 (Promega Corporation).

2. The plasmid pSSPGuyHuA2 is 5296 bp in length. Nucleotides 49-1165 represent the Superpromoter (Ni et al., 1995). Nucleotides 1166-2645 (Figure 6) comprise a sequence encoding a mouse heavy chain variable/human IgA2m(2) constant hybrid with linker sequences. A consensus Kozak sequence (Kozak, 1986) is included (nt 1186-1192) to enhance translation initiation, and the native mouse signal peptide has been replaced with the signal peptide from bean legumin (nt 1189-1257; Bäumlein et al., 1986). The variable region (nt 1258-1595) is from the murine monoclonal Guy's 13 (Smith and Lehner, 1989). The sequence of the human IgA2m(2) constant region (nt 1596-2645) has been previously published (Chintalacharuvu, et al., 1994). Nucleotides 2646-2371 derive from the nopaline synthase 3' end (transcription termination and polyadenlyation signal; Depicker et al., 1982). The remainder of the plasmid derives from the vector pSP72 (Promega).

3. The plasmid pSHuJ is 4283 bp in length. Nucleotides 14-1136 represent the Superpromoter (Ni et al., 1995) and nucleotides 1137-1648 are shown in Figure 7 and comprise a sequence encoding the human J chain including the native signal peptide (Max et al, 1985) along with linker sequences. A consensus Kozak sequence (Kozak, 1986) is included (nt 1162-1168) to enhance translation initiation. Nucleotides 1649-1902 derive from the nopaline synthase 3' end (transcription termination and polyadenlyation signal; Depicker et al., 1982). The remainder of the plasmid derives from the vector pSP72 (Promega).

4. The plasmid pSHuSC is 5650 bp in length. Nucleotides 13-1136 are derived from the Superpromoter (Ni et al., 1995) and nucleotides 1137-2981 are shown in Figure 8 and comprise a sequence encoding the human Secretory Component including the native signal peptide (Krajci et al., 1989) along with linker sequences. A consensus Kozak sequence (Kozak, 1986) is included (nt 1151-1157) to enhance translation initiation. Nucleotides 2982-3236 derive from the nopaline synthase 3' end (transcription termination and polyadenlyation

signal; Depicker et al., 1982). The remainder of the plasmid derives from the vector pSP72 (Promega).

Particle Bombardment and Regeneration of Plant

The four vectors described above, containing heavy chain, light chain, J chain and secretory component, were used for particle bombardment of tobacco leaf disks using Biolistic® PDS-1000 HE instrument (Bio-Rad). A fifth vector containing an antibiotic resistance gene, such as pSZeo (containing the Phe^r gene encoding resistance to the antibiotic Zeocin; Perez et al., 1989; Drocourt et al, 1990) or pBMSP-1 (containing the Kan^r gene encoding resistance to kanamycin; S. Gelvin, Purdue University) was also used. Transgenic plants were screened using chain-specific antibodies by Western blot to identify individual transformants expressing assembled human SIgA.

A stock suspension of microprojectiles was prepared by mixing 60 mg of 0.7 micron tungsten or gold particles in 1 mL of 70% ethanol in a sterile microcentrifuge tube. This suspension was vortexed 5 minutes and incubated at room temperature for 5 minutes. After microcentrifuging for 5 seconds the ethanol was removed and the pellet was resuspended in 1ml sterile water and centrifuged for 5 minutes. Particles were washed 3 times with sterile water, removing wash each time after a brief centrifugation. Sterile 50% glycerol was then added to particles to bring the concentration to 60 mg/mL. This suspension was dispensed in 100 microliter aliquots in sterile 1.5 mL microcentrifuge tubes.

An aliquot of 60 mg of particles (in 50% glycerol) was vortexed for 5 minutes. While vortexing, 50 microliters of this suspension was removed to a sterile 0.5 mL microfuge tube. While vortexing this tube, 1 microgram of each plasmid DNA was added. While vortexing, 50 ml of CaCl2 (2.5M), and 20 ml of spermidine (0.1 M) were added (in that order). This mixture was vortexed for another 3 minutes, allowed to settle for 2 min, and then centrifuged for 5 seconds. The supernatant was removed and 140 ml of absolute ethanol was added to the DNA coated particles. The particles were allowed to settle for 5 minutes, and the supernatant was removed. The particles were resuspended in 140 ml of absolute

ethanol and allowed to settle a second time. The supernatant was removed, and the particles were resuspended in 50 ml of absolute ethanol.

Ten ml of the suspension was applied as evenly as possible onto the center of macrocarrier sheet made of Kapton (DuPont) and the ethanol was evaporated. The macrocarrier sheet and a rupture disk were placed in the unit (Bio-Rad). A petri dish containing pieces of a surface-sterilized tobacco leaf was placed below the stopping screen. The chamber was evacuated to 28-29mm Hg and the target was bombarded once.

Tobacco leaf disks were bombarded with tungsten particles (1 μm) coated with the plasmids pSSpGuyHuK, pSSpGuyHuA2, pSHuJ, pSHuSC and pBMSP-1. Leaves from plants grown axenically were used, with the optimal transformation efficiency resulting from leaves taken 10 days after transplanting a node to new media. Regeneration of bombarded leaf disks was performed essentially according to standard protocols, for example Horsch et al., 1985. Leaf disks were immediately placed on MS media supplemented with NAA and BAP, along with 50 mg/L kanamycin sulfate. After 4 days, leaves were cut into small segments, dipped in water and placed on fresh plates. Leaves were dipped in water every week. Small calli appeared between 1 and 2 months after bombardment, and these were maintained under selection until shoots appeared about 2 weeks later. Shoots were transferred to fresh media, and when leaves were at least 3 cm in length, cuttings were taken and extracted for ELISA and immunoblot analysis. At least 1 of 10 putative transgenic plants produced fully assembled SIgA.

Screening of transgenic plants - Western blot analysis of immunoglobulin

We have found that immunoblotting is a sensitive and reliable method to detect expression of antibody chains in plants. Two 4 mm diameter leaf punches were homogenized with 75 mM Tris-HCl (pH 6.8), 2% SDS, under reducing and non-reducing conditions (± DTT). Homogenates were centrifuged to pellet debris and supernatants were loaded onto SDS poly-acrylamide gels. Electrophoresis

was performed and the gels were blotted onto nitrocellulose. Blots were rinsed 2 times for 10 min in TBST (TBS with 0.05% Tween 20), then incubated for 1 hr in TBST + 5% non-fat dry milk. Blots were rinsed twice again with TBST, then incubated for 1 hr at room temperature with titered goat anti-human chain-specific HRP-conjugated antisera in TBST + 2% goat serum. After washing twice with TBST and once with TBS, antibody binding was detected by incubation with chemiluminescent reagents (Pierce), and exposure of the blot to x-ray film.

Example 2 Transformation of Duckweed with SIgA

In a transformation system for *Lemna gibba*, bombardment of surface-sterilized leaf tissue with DNA-coated particles is much the same as with tobacco. Whole thalloid fronds and excised portions of the frond having meristematic areas of the plus and minus pockets are used. Surface sterilization is accomplished by treating with a solution of sodium dichloriisocyanurate for 20-25 min followed by three rinses in sterile water. Subsequent to bombardment, the explants are placed on modified medium of Hillman (1961), having the following constituents:

Ingredient	Concentration (mg/L)
KH_2PO_4	680
KNO_3	1515
Ca(NO3)2 • 4H20	1180
MgSO4 •7H2O	500
H3BO3	2.86
ZnSO4 •7H2O	0.22
Na2MoO4•2H2O	0.12
CuSO4•5H2O	0.08
MnCl2•4H2O	3.62
FeCl3•6H2O	5.4
Tartaric acid	3.00
EDTA	9.0
Sucrose	4,000-10,000
Coconut milk	10% (v/v)

Also included in the medium are a kinetin, zeatin riboside, thidiazuron and 2,4-D, separately or in combinations (concentration ranging from 0.1 to 10 mg/L). Culture conditions are 28±1°C at a relative humidity of 65-75% and a day length

of 10 hours. Light is provided by day light fluorescent tubes. After 5 days the leaf discs are transferred to regeneration medium containing Zeocin to prevent regeneration from untransformed plant cells. After three weeks, the callus that forms from these leaf disks are transferred to medium without hormones to stimulate the development of somatic embryos. After 3-4 months, regenerated plants are then screened for production of immunoglobulin chains and assembled SIgA.

Explants are incubated in the light under these conditions until whole new fronds and roots form, after which the new plants are transferred to liquid medium and allowed to propagate clonally. Conditions for maintaining *Lemna* in aquaculture are described by Porath (U.S.Patent # 5,269,819) which is herein incorporated by reference. When sufficient material is available, it is harvested and extracted for ELISA and immunoblot analysis. We expect to find fully assembled SIgA.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for producing a multimeric protein in a plant cell wherein the multimeric protein is heterologous to the plant cell, the method comprising the steps of:

(a) transforming a plant cell with a plurality of naked plasmids, each plasmid encoding less than all of the polypeptide components of the multimeric protein, and said plurality encoding all of the polypeptide components of the multimeric protein; and

(b) culturing the plant cell under conditions suitable for protein expression, thereby producing the multimeric protein.

2. The method of claim 1, further comprising the step of isolating the produced multimeric protein from the cell.

3. The method of claim 1, wherein the plant cell is intact.

4. The method of claim 1, wherein the multimeric protein is biologically active.

5. The method of claim 1, wherein each plasmid encodes a single polypeptide component of the multimeric protein.

6. The method of claim 1, wherein at least one plasmid encodes multiple polypeptide components of the multimeric protein.

7. The method of claim 1, wherein at least one plasmid comprises a sequence encoding a signal peptide.

8. The method of claim 1, wherein at least one plasmid comprises a sequence encoding the amino acid sequence KDEL.

9. The method of claim 1, wherein at least one plasmid comprises a sequence encoding a selectable marker.

10. The method of claim 1, wherein the plant cell is from a dicotyledonous plant.

11. The method of claim 1, wherein the plant cell is from a monocotylendous plant.

12. The method of claim 10, wherein said dicotyledonous plant is tobacco.

13. The method of claim 11, wherein said monotyledonous plant is *Lemna gibba* (L.)

14. The method of claim 1, wherein the multimeric protein is selected from the group consisting of an immunoglobulin molecule, a receptor-ligand complex, a receptor homodimer, a receptor herterodimer, and a trimeric G-protein.

15. The method of claim 14, wherein the immunoglobulin molecule is selected from the group consisting of IgA, IgM, IgG, IgD, and IgE.

16. The method of claim 14, wherein the immunoglobulin molecule is IgA.

17. Microparticles coated with a plurality of plasmids, each plasmid encoding less than all of the polypeptide components of a multimeric protein, and

said plurality encoding all of the polypeptide components of the multimeric protein.

18. The microparticles of claim 17, wherein the microparticles are tungsten or gold.

19. A transgenic plant or plant cell expressing a multimeric protein that is heterologous to the plant cell, wherein said plants or plant cells are characterized by adjacent integration of multiple expression cassettes, each expression cassette encoding less than all of the polypeptide components of the multimeric protein, and said multiple expression cassettes encoding all of the polypeptide components of the multimeric protein.

20. The method of claim 16, wherein the IgA molecule is secretory.

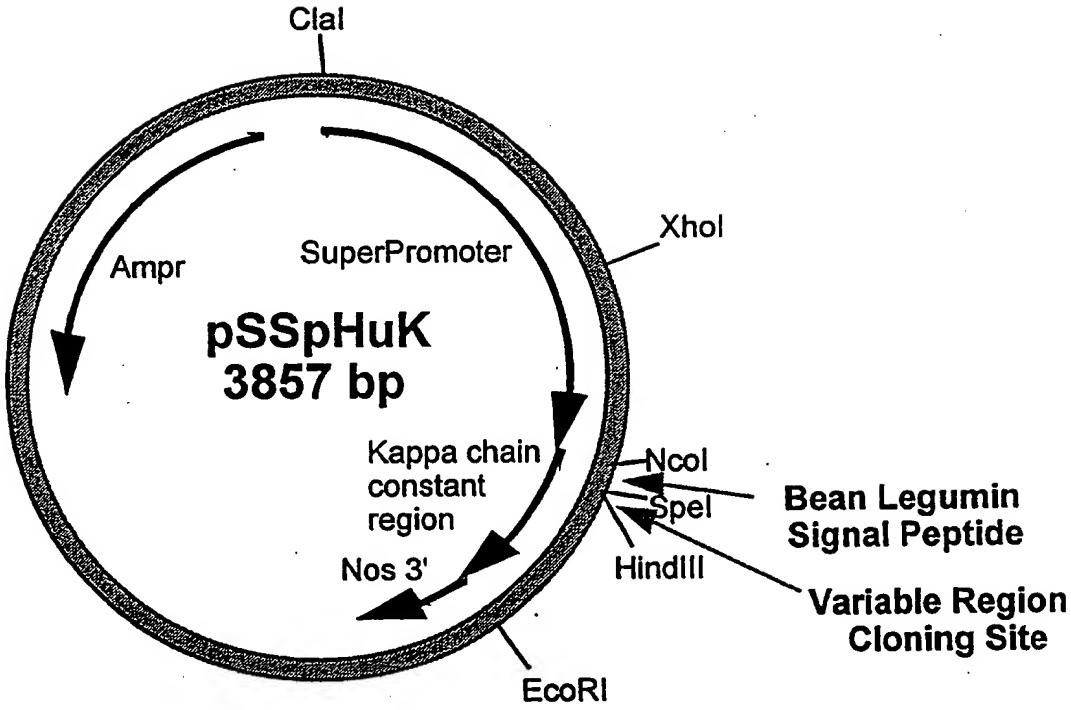
ClaI
XhoI
Ampr
SuperPromoter
pSSpHuK
3857 bp
Kappa chain
constant
region
NcoI
SpeI
HindIII
Nos 3'
EcoRI
Bean Legumin
Signal Peptide
Variable Region
Cloning Site

Figure 1

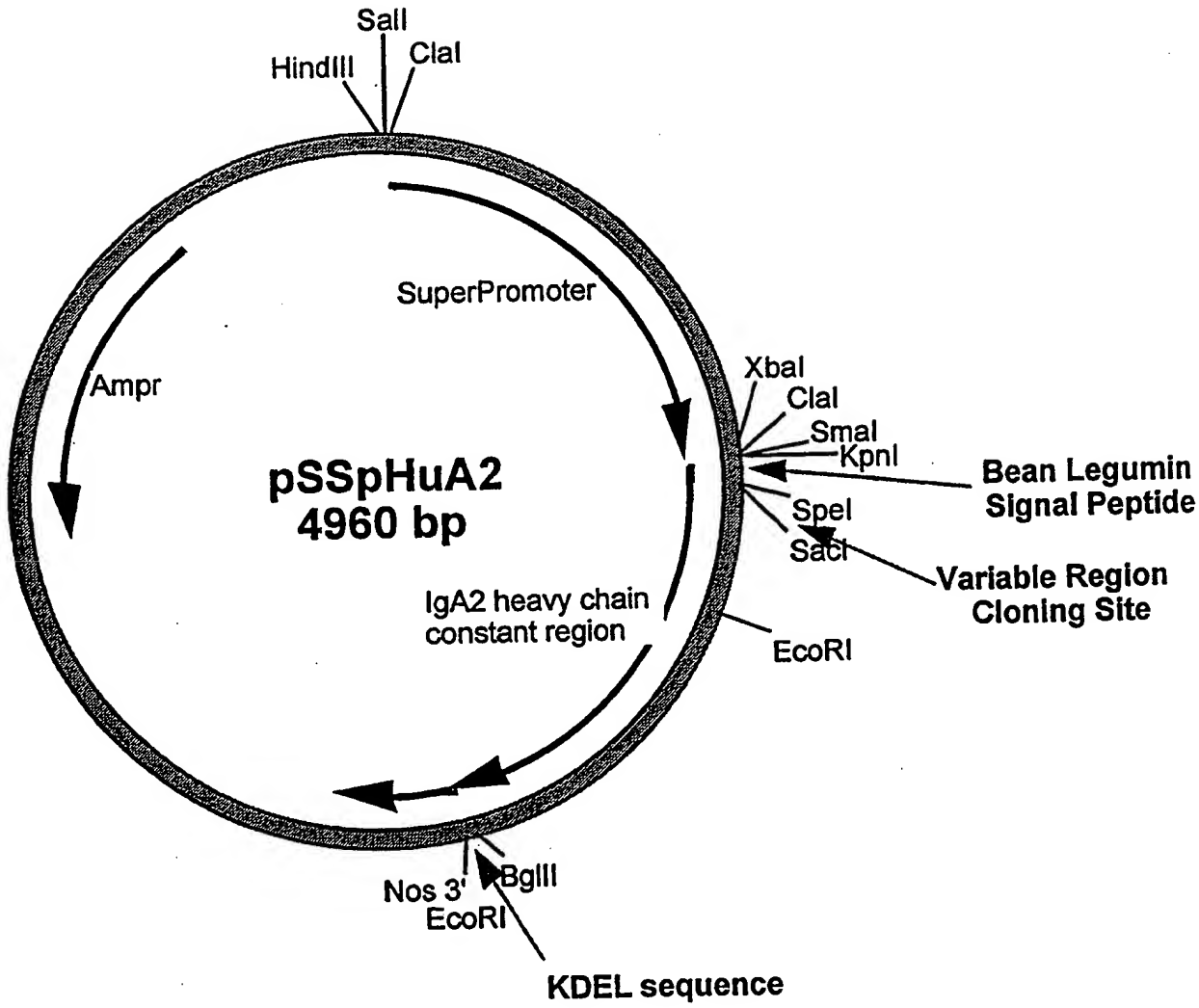
SalI
HindIII
ClaI
SuperPromoter
Ampr
XbaI
ClaI
SmaI
KpnI
Bean Legumin
Signal Peptide
SpeI
SacI
Variable Region
Cloning Site
pSSpHuA2
4960 bp
IgA2 heavy chain
constant region
EcoRI
BglII
Nos 3'
EcoRI
KDEL sequence

Figure 2

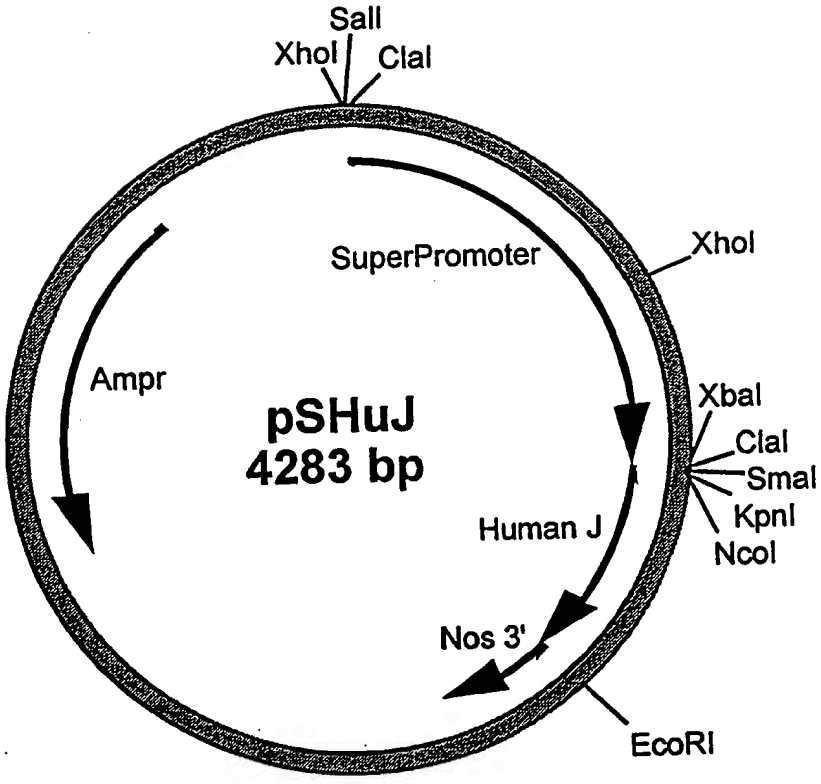
SalI
XhoI
ClaI
SuperPromoter
XhoI
Ampr
pSHuJ
4283 bp
XbaI
ClaI
SmaI
KpnI
NcoI
Human J
Nos 3'
EcoRI

Figure 3

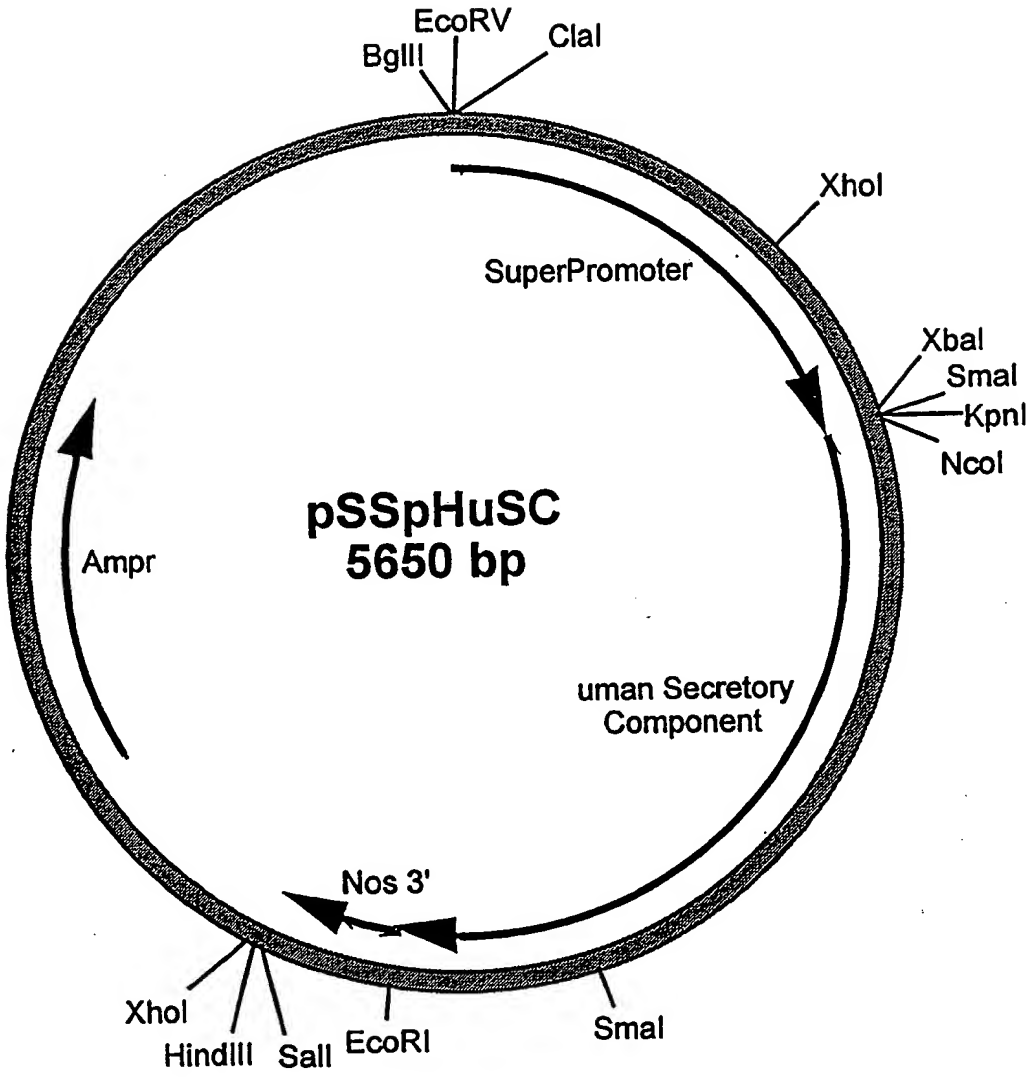
EcoRV
ClaI
BglII
XhoI
SuperPromoter
XbaI
SmaI
KpnI
NcoI
pSSpHuSC
5650 bp
Ampr
uman Secretory
Component
Nos 3'
XhoI
HindIII
SalI
EcoRI
SmaI

Figure 4

Figure 5

```
1140
 GGATCTAACC ATG GGA TCT AAA CCT TTT TTG TCT CTT CTT TCA TTG TCA TTG CTT TTG TTT
            met gly ser lys pro phe leu ser leu leu ser leu ser leu leu leu phe
1201/18  Spe I
ACA TCT ACT AGT TTG GCA GAC ATT GTG ATG ACC CAG TCT CCA GCA ATC ATG TCT GCA TCT
thr ser thr ser leu ala asp ile val met thr gln ser pro ala ile met ser ala ser
1261/38
CCA GGG GAG AAG GTC ACC ATA ACC TGC AGT GCC AGC TCA AGT GTA AGT TAC ATG CAC TGG
pro gly glu lys val thr ile thr cys ser ala ser ser ser val ser tyr met his trp
1321/58
TTC CAG CAG AAG CCA GGC ACT TCT CCC AAA CTC TGG CTT TAT AGC ACA TCC AAC CTG GCT
phe gln gln lys pro gly thr ser pro lys leu trp leu tyr ser thr ser asn leu ala
1381/78
TCT GGA GTC CCT GCT CGC TTC AGT GGC AGT GGA TCT GGG ACC TCT TAC TCT CTC ACA ATC
ser gly val pro ala arg phe ser gly ser gly ser gly thr ser tyr ser leu thr ile
1441/98
AGC CGA ATG GAG GCT GAA GAT GCT GCC ACT TAT TAC TGC CAT CAA AGG ACT AGC TAC CCA
ser arg met glu ala glu asp ala ala thr tyr tyr cys his gln arg thr ser tyr pro
1501/118                    Hind III
TAC ACA TTC GGA GGG GGG ACC AAG CTT GAG ATC AAA CGA ACT GTG GCT GCA CCA TCT GTC
tyr thr phe gly gly gly thr lys leu glu ile lys arg thr val ala ala pro ser val
1561/138
TTC ATC TTC CCG CCA TCT GAT GAG CAG TTG AAA TCT GGA ACT GCC TCT GTT GTG TGC CTG
phe ile phe pro pro ser asp glu gln leu lys ser gly thr ala ser val val cys leu
1621/158
CTG AAT AAC TTC TAT CCC AGA GAG GCC AAA GTA CAG TGG AAG GTG GAT AAC GCC CTC CAA
leu asn asn phe tyr pro arg glu ala lys val gln trp lys val asp asn ala leu gln
1681/178
TCG GGT AAC TCC CAG GAG AGT GTC ACA GAG CAG GAC AGC AAG GAC AGC ACC TAC AGC CTC
ser gly asn ser gln glu ser val thr glu gln asp ser lys asp ser thr tyr ser leu
1741/198
AGC AGC ACC CTG ACG CTG AGC AAA GCA GAC TAC GAG AAA CAC AAA GTC TAC GCC TGC GAA
ser ser thr leu thr leu ser lys ala asp tyr glu lys his lys val tyr ala cys glu
1801/218
GTC ACC CAT CAG GGC CTG AGC TCG CCC GTC ACA AAG AGC TTC AAC AGG GGA GAG TGT TGA
val thr his gln gly leu ser ser pro val thr lys ser phe asn arg gly glu cys
1861/237
ATTC
```

Figure 6

```
1166
                                                GGATCTATCGATTCCCGGGTACC ATG GGA TCT AAA
                                                                        met gly ser lys
1201/5                                                                  Spe I
CCT TTT TTG TCT CTT CTT TCA TTG TCA TTG CTT TTG TTT ACA TCT ACT AGT TTG GCA GGG
pro phe leu ser leu leu ser leu ser leu leu leu phe thr ser thr ser leu ala gly
1261/25
GTC CAG CTT CAG CAG TCA GGA CCT GAC CTG GTG AAA CCT GGG GCC TCA GTG AAG ATA TCC
val gln leu gln gln ser gly pro asp leu val lys pro gly ala ser val lys ile ser
1321/45
TGC AAG GCT TCT GGA TAC ACA TTC ACT GAC TAC AAC ATA CAC TGG GTG AAG CAG AGC CGT
cys lys ala ser gly tyr thr phe thr asp tyr asn ile his trp val lys gln ser arg
1381/65
GGA AAG AGC CTT GAG TGG ATT GGA TAT ATT TAT CCT TAC AAT GGT AAT ACT TAC TAC AAC
gly lys ser leu glu trp ile gly tyr ile tyr pro tyr asn gly asn thr tyr tyr asn
1441/85
CAG AAG TTC AAG AAC AAG GCC ACA TTG ACT GTA GAC AAT TCC TCC ACC TCA GCC TAC ATG
gln lys phe lys asn lys ala thr leu thr val asp asn ser ser thr ser ala tyr met
1501/105
GAG CTC CGC AGC CTG ACA TCT GAG GAC TCT GCA GTC TAT TAC TGT GCA ACC TAC TTT GAC
glu leu arg ser leu thr ser glu asp ser ala val tyr tyr cys ala thr tyr phe asp
1561/125                                Sac I
TAC TGG GGC CAA GGC ACC ACT CTC ACA GTG AGC TCA GCA TCC CCG ACC AGC CCC AAG GTC
tyr trp gly gln gly thr thr leu thr val ser ser ala ser pro thr ser pro lys val
1621/145
TTC CCG CTG AGC CTC GAC AGC ACC CCC CAA GAT GGG AAC GTG GTC GTC GCA TGC CTG GTC
phe pro leu ser leu asp ser thr pro gln asp gly asn val val val ala cys leu val
1681/165
CAG GGC TTC TTC CCC CAG GAG CCA CTC AGT GTG ACC TGG AGC GAA AGC GGA CAG AAC GTG
gln gly phe phe pro gln glu pro leu ser val thr trp ser glu ser gly gln asn val
1741/185
ACC GCC AGA AAC TTC CCA CCT AGC CAG GAT GCC TCC GGG GAC CTG TAC ACC ACG AGC AGC
thr ala arg asn phe pro pro ser gln asp ala ser gly asp leu tyr thr thr ser ser
 pro pro glu thr ser his leu ala arg met pro pro gly thr cys thr pro arg ala ala
  arg gln lys leu pro thr AMB pro gly cys leu arg gly pro val his his glu gln pro
1801/205
CAG CTG ACC CTG CCG GCC ACA CAG TGC CCA GAC GGC AAG TCC GTG ACA TGC CAC GTG AAG
gln leu thr leu pro ala thr gln cys pro asp gly lys ser val thr cys his val lys
1861/225
CAC TAC ACG AAT TCC AGC CAG GAT GTG ACT GTG CCC TGC CGA GTT CCC CCA CCT CCC CCA
his tyr thr asn ser ser gln asp val thr val pro cys arg val pro pro pro pro pro
1921/245
TGC TGC CAC CCC CGA CTG TCG CTG CAC CGA CCG GCC CTC GAG GAC CTG CTC TTA GGT TCA
cys cys his pro arg leu ser leu his arg pro ala leu glu asp leu leu leu gly ser
1981/265
GAA GCG AAC CTC ACG TGC ACA CTG ACC GGC CTG AGA GAT GCC TCT GGT GCC ACC TTC ACC
glu ala asn leu thr cys thr leu thr gly leu arg asp ala ser gly ala thr phe thr
2041/285
TGG ACG CCC TCA AGT GGG AAG AGC GCT GTT CAA GGA CCA CCT GAG CGT GAC CTC TGT GGC
trp thr pro ser ser gly lys ser ala val gln gly pro pro glu arg asp leu cys gly
```

2101/305
TGC TAC AGC GTG TCC AGA GTA CTT CCT GGC TGT GCC CAG CCA TGG AAC CAT GGG GAG ACC
cys tyr ser val ser arg val leu pro gly cys ala gln pro trp asn his gly glu thr
2161/325
TTC ACC TGC ACT GCT GCC CAC CCC GAG TTG AAG ACC CCA CTA ACC GCC AAC ATC ACA AAA
phe thr cys thr ala ala his pro glu leu lys thr pro leu thr ala asn ile thr lys
2221/345
TCC GGA AAC ACA TTC CGG CCC GAG GTC CAC CTG CTG CCG CCG CCG TCG GAG GAG CTG GCC
ser gly asn thr phe arg pro glu val his leu leu pro pro pro ser glu glu leu ala
2281/365
CTG AAC GAG CTG GTG ACG CTG ACG TGC CTG GCA CGT GGC TTC AGC CCC AAG GAT GTG CTG
leu asn glu leu val thr leu thr cys leu ala arg gly phe ser pro lys asp val leu
2341/385
GTT CGC TGG CTG CAG GGG TCA CAG GAG CTG CCC CGC GAG AAG TAC CTG ACT TGG GCA TCC
val arg trp leu gln gly ser gln glu leu pro arg glu lys tyr leu thr trp ala ser
2401/405
CGG CAG GAG CCC AGC CAG GGC ACC ACC ACC TAT GCT GTG ACC AGC ATA CTG CGC GTG GCA
arg gln glu pro ser gln gly thr thr thr tyr ala val thr ser ile leu arg val ala
2461/425
GCC GAG GAC TGG AAG AAG GGG GAG ACC TTC TCC TGC ATG GTG GGC CAC GAG GCC CTG CCG
ala glu asp trp lys lys gly glu thr phe ser cys met val gly his glu ala leu pro
2521/445
CTG GCC TTC ACA CAG AAG ACC ATC GAC CGC TTG GCG GGT AAA CCC ACC CAT ATC AAT GTG
leu ala phe thr gln lys thr ile asp arg leu ala gly lys pro thr his ile asn val
2581/465
TCT GTT GTC ATG GCG GAG GCG GAC GGC ACC TGC TAC AGA TCT GAA AAG GAT GAA CTT TAG
ser val val met ala glu ala asp gly thr cys tyr arg ser glu lys asp glu leu
2641/483
AAT TC

Figure 7

```
1141
      AGGATCTATCGATTCCCGGGTACC ATG GAG AAC CAT TTG CTT TTC TGG GGA GTC CTG GCG
                               met glu asn his leu leu phe trp gly val leu ala
1201/13
GTT TTT ATT AAG GCT GTT CAT GTG AAA GCC CAA GAA GAT GAA AGG ATT GTT CTT GTT GAC
val phe ile lys ala val his val lys ala gln glu asp glu arg ile val leu val asp
1261/33
AAC AAA TGT AAG TGT GCC CGG ATT ACT TCC AGG ATC ATC CGT TCT TCC GAA GAT CCT AAT
asn lys cys lys cys ala arg ile thr ser arg ile ile arg ser ser glu asp pro asn
1321/53
GAG GAC ATT GTG GAG AGA AAC ATC CGA ATT ATT GTT CCT CTG AAC AAC AGG GAG AAT ATC
glu asp ile val glu arg asn ile arg ile ile val pro leu asn asn arg glu asn ile
1381/73
TCT GAT CCC ACC TCA CCA TTG AGA ACC AGA TTT GTG TAC CAT TTG TCT GAC CTC TGT AAA
ser asp pro thr ser pro leu arg thr arg phe val tyr his leu ser asp leu cys lys
1441/93
AAA TGT GAT CCT ACA GAA GTG GAG CTG GAT AAT CAG ATA GTT ACT GCT ACC CAG AGC AAT
lys cys asp pro thr glu val glu leu asp asn gln ile val thr ala thr gln ser asn
1501/113
ATC TGT GAT GAA GAC AGT GCT ACA GAG ACC TGC TAC ACT TAT GAC AGA AAC AAG TGC TAC
ile cys asp glu asp ser ala thr glu thr cys tyr thr tyr asp arg asn lys cys tyr
1561/133
ACA GCT GTG GTC CCA CTC GTA TAT GGT GGT GAG ACC AAA ATG GTG GAA ACA GCC TTA ACC
thr ala val val pro leu val tyr gly gly glu thr lys met val glu thr ala leu thr
1621/153
CCA GAT GCC TGC TAT CCT GAC TGA ATTC
pro asp ala cys tyr pro asp
```

Figure 8

```
1137
                                    GTCGATTCCCGGGTACC ATG GTG CTC TTC GTG CTC ACC TGC
                                                      met val leu phe val leu thr cys
1178/9
CTG CTG GCG GTC TTC CCA GCC ATC TCC ACG AAG AGT CCC ATA TTT GGT CCC GAG GAG GTG
leu leu ala val phe pro ala ile ser thr lys ser pro ile phe gly pro glu glu val
1238/29
AAT AGT GTG GAA GGT AAC TCA GTG TCC ATC ACG TGC TAC TAC CCA CCC ACC TCT GTC AAC
asn ser val glu gly asn ser val ser ile thr cys tyr tyr pro pro thr ser val asn
1298/49
CGG CAC ACC CGG AAG TAC TGG TGC CGG CAG GGA GCT AGA GGT GGC TGC ATA ACC CTC ATC
arg his thr arg lys tyr trp cys arg gln gly ala arg gly gly cys ile thr leu ile
1358/69
TCC TCG GAG GGC TAC GTC TCC AGC AAA TAT GCA GGC AGG GCT AAC CTC ACC AAC TTC CCG
ser ser glu gly tyr val ser ser lys tyr ala gly arg ala asn leu thr asn phe pro
1418/89
GAG AAC GGC ACA TTT GTG GTG AAC ATT GCC CAG CTG AGC CAG GAT GAC TCC GGG CGC TAC
glu asn gly thr phe val val asn ile ala gln leu ser gln asp asp ser gly arg tyr
1478/109
AAG TGT GGC CTG GGC ATC AAT AGC CGA GGC CTG TCC TTT GAT GTC AGC CTG GAG GTC AGC
lys cys gly leu gly ile asn ser arg gly leu ser phe asp val ser leu glu val ser
1538/129
CAG GGT CCT GGG CTC CTA AAT GAC ACT AAA GTC TAC ACA GTG GAC CTG GGC AGA ACG GTG
gln gly pro gly leu leu asn asp thr lys val tyr thr val asp leu gly arg thr val
1598/149
ACC ATC AAC TGC CCT TTC AAG ACT GAG AAT GCT CAA AAG AGG AAG TCC TTG TAC AAG CAG
thr ile asn cys pro phe lys thr glu asn ala gln lys arg lys ser leu tyr lys gln
1658/169
ATA GGC CTG TAC CCT GTG CTG GTC ATC GAC TCC AGT GGT TAT GTG AAT CCC AAC TAT ACA
ile gly leu tyr pro val leu val ile asp ser ser gly tyr val asn pro asn tyr thr
1718/189
GGA AGA ATA CGC CTT GAT ATT CAG GGT ACT GGC CAG TTA CTG TTC AGC GTT GTC ATC AAC
gly arg ile arg leu asp ile gln gly thr gly gln leu leu phe ser val val ile asn
1778/209
CAA CTC AGG CTC AGC GAT GCT GGG CAG TAT CTC TGC CAG GCT GGG GAT GAT TCC AAT AGT
gln leu arg leu ser asp ala gly gln tyr leu cys gln ala gly asp asp ser asn ser
1838/229
AAT AAG AAG AAT GCT GAC CTC CAA GTG CTA AAG CCC GAG CCC GAG CTG GTT TAT GAA GAC
asn lys lys asn ala asp leu gln val leu lys pro glu pro glu leu val tyr glu asp
1898/249
CTG AGG GGC TCA GTG ACC TTC CAC TGT GCC CTG GGC CCT GAG GTG GCA AAC GTG GCC AAA
leu arg gly ser val thr phe his cys ala leu gly pro glu val ala asn val ala lys
1958/269
TTT CTG TGC CGA CAG AGC AGT GGG GAA AAC TGT GAC GTG GTC GTC AAC ACC CTG GGG AAG
phe leu cys arg gln ser ser gly glu asn cys asp val val val asn thr leu gly lys
2018/289
AGG GCC CCA GCC TTT GAG GGC AGG ATC CTG CTC AAC CCC CAG GAC AAG GAT GGC TCA TTC
arg ala pro ala phe glu gly arg ile leu leu asn pro gln asp lys asp gly ser phe
```

```
2078/309
AGT GTG GTG ATC ACA GGC CTG AGG AAG GAG GAT GCA GGG CGC TAC CTG TGT GGA GCC CAT
ser val val ile thr gly leu arg lys glu asp ala gly arg tyr leu cys gly ala his
2138/329
TCG GAT GGT CAG CTG CAG GAA GGC TCG CCT ATC CAG GCC TGG CAA CTC TTC GTC AAT GAG
ser asp gly gln leu gln glu gly ser pro ile gln ala trp gln leu phe val asn glu
2198/349
GAG TCC ACG ATT CCC CGC AGC CCC ACT GTG GTG AAG GGG GTG GCA GGA AGC TCT GTG GCC
glu ser thr ile pro arg ser pro thr val val lys gly val ala gly ser ser val ala
2258/369
GTG CTC TGC CCC TAC AAC CGT AAG GAA AGC AAA AGC ATC AAG TAC TGG TGT CTC TGG GAA
val leu cys pro tyr asn arg lys glu ser lys ser ile lys tyr trp cys leu trp glu
2318/389
GGG GCC CAG AAT GGC CGC TGC CCC CTG CTG GTG GAC AGC GAG GGG TGG GTT AAG GCC CAG
gly ala gln asn gly arg cys pro leu leu val asp ser glu gly trp val lys ala gln
2378/409
TAC GAG GGC CGC CTC TCC CTG CTG GAG GAG CCA GGC AAC GGC ACC TTC ACT GTC ATC CTC
tyr glu gly arg leu ser leu leu glu glu pro gly asn gly thr phe thr val ile leu
2438/429
AAC CAG CTC ACC AGC CGG GAC GCC GGC TTC TAC TGG TGT CTG ACC AAC GGC GAT ACT CTC
asn gln leu thr ser arg asp ala gly phe tyr trp cys leu thr asn gly asp thr leu
2498/449
TGG AGG ACC ACC GTG GAG ATC AAG ATT ATC GAA GGA GAA CCA AAC CTC AAG GTT CCC GGG
trp arg thr thr val glu ile lys ile ile glu gly glu pro asn leu lys val pro gly
2558/469
AAT GTC ACG GCT GTG CTG GGA GAG ACT CTC AAG GTC CCC TGT CAC TTT CCA TGC AAA TTC
asn val thr ala val leu gly glu thr leu lys val pro cys his phe pro cys lys phe
2618/489
TCC TCG TAC GAG AAA TAC TGG TGC AAG TGG AAT AAC ACG GGC TGC CAG GCC CTG CCC AGC
ser ser tyr glu lys tyr trp cys lys trp asn asn thr gly cys gln ala leu pro ser
2678/509
CAA GAC GAA GGC CCC AGC AAG GCC TTC GTG AAC TGT GAC GAG AAC AGC CGG CTT GTC TCC
gln asp glu gly pro ser lys ala phe val asn cys asp glu asn ser arg leu val ser
2738/529
CTG ACC CTG AAC CTG GTG ACC AGG GCT GAT GAG GGC TGG TAC TGG TGT GGA GTG AAG CAG
leu thr leu asn leu val thr arg ala asp glu gly trp tyr trp cys gly val lys gln
2798/549
GGC CAC TTC TAT GGA GAG ACT GCA GCC GTC TAT GTG GCA GTT GAA GAG AGG AAG GCA GCG
gly his phe tyr gly glu thr ala ala val tyr val ala val glu glu arg lys ala ala
2858/569
GGG TCC CGC GAT GTC AGC CTA GCG AAG GCA GAC GCT GCT CCT GAT GAG AAG GTG CTA GAC
gly ser arg asp val ser leu ala lys ala asp ala ala pro asp glu lys val leu asp
2918/589
TCT GGT TTT CGG GAG ATT GAG AAC AAA GCC ATT CAG GAT CCC AGG CTT TTT GCA GAG TGA
ser gly phe arg glu ile glu asn lys ala ile gln asp pro arg leu phe ala glu
2978
ATTC
```